

CLAIMS

What is claimed is:

1 1. A machine-implemented method for extrapolating user profile
2 information from web page access patterns of a user, comprising:
3 detecting a set of web pages accessed by a user;
4 mapping at least a subset of said web pages to a first data structure, said
5 first data structure identifies a web page access pattern of said user;
6 comparing said first data structure to a second data structure to obtain a
7 comparison result, said second data structure identifies a web page access pattern
8 of a set of known users, said known users having a user profile attribute in
9 common; and
10 assigning said user profile attribute to said user in response to said
11 comparison result.

1 2. The method of claim 1, wherein said first and second data
2 structures are multi-dimensional vectors, wherein each dimension of said vectors
3 corresponds to a web page.

1 3. The method of claim 2, wherein said comparing step comprises:
2 determining a distance between said vectors.

1 4. The method of claim 2, wherein said comparison result is a cosine
2 of an angle between said vectors.

1 5. The method of claim 1, wherein said profile attribute is
2 demographic information.

1 6. The method of claim 1, wherein said profile attribute is a gender of
2 said user.

1 7. A machine-implemented method for extrapolating user profile
2 information from web page access patterns of a user, comprising:
3 assigning bias values to a plurality of web pages;
4 detecting at least a subset of said web pages accessed by a user;
5 combining said bias values of said subset of web pages to obtain a
6 combination result; and
7 assigning a user profile attribute to said user in response to said
8 combination result.

1 8. The method of claim 7, wherein said combination result is a
2 summation of said bias values of said subset of web pages.

1 9. The method of claim 7, wherein said profile attribute is
2 demographic information.

1 10. The method of claim 7, wherein said profile attribute is a gender of
2 said user.

1 11. A machine-implemented method for extrapolating profile
2 information from web page access patterns of a test user, comprising:
3 detecting a set of web pages accessed by a test user;
4 initializing a first set of Expectation Maximization (EM) parameters;
5 performing a first EM process using said first set of initialized parameters
6 to obtain a first EM process result; and
7 assigning a user profile attribute to said test user in response to said first
8 EM process result.

1 12. The method of claim 11, wherein said EM process comprises the
2 steps of:
3 performing a first expectation step using said first set of initialized
4 parameters to obtain an expectation result;
5 performing a first maximization step using said expectation result to
6 obtain a maximization result; and
7 repeating said expectation and maximization steps, said repeated
8 expectation step uses said maximization result.

1 13. The method of claim 12, further comprising:
2 determining a log-likelihood in place of said repeating step; and
3 repeating said expectation and maximization steps in response to said log-
4 likelihood.

1 14. The method of claim 12, further comprising:

2 determining an accuracy value on a separate validation data set of said
3 repeating step; and
4 repeating said expectation and maximization steps in response to said
5 accuracy value.

1 15. The method of claim 12, wherein said expectation result is a
2 conditional probability of a gender given a web page and a user.

1 16. The method of claim 12, wherein said maximization result is a
2 conditional probability of a web page given a gender.

1 17. The method of claim 12, wherein said maximization result is a
2 conditional probability of a gender given a user.

1 18. The method of claim 11, wherein:
2 said first set of EM parameters comprises data from a training set of users,
3 said method further comprising:
4 initializing a second set of EM parameters in place of said
5 assigning step, said second set of EM parameters comprise web page access
6 information for said test user;
7 performing a second EM process using said first EM process result
8 and said second set of EM parameters to obtain a second EM process result; and
9 assigning a user profile attribute to said test user in response to
10 said second EM process result.

1 19. The method of claim 18, wherein at least one profile attribute of
2 said users of said training set is known prior to the performance of said first EM
3 process.

1 20. The method of claim 11, wherein said first EM process is
2 performed on a set of users, wherein at least one profile attribute of at least a
3 subset of said set of users is known prior to the performance of said first EM
4 process.

1 21. The method of claim 11, wherein said first EM process result is a
2 probability of said user profile attribute given said test user.

1 22. The method of claim 21, wherein said user profile attribute is
2 assigned to said test user if said probability is greater than a threshold value.

1 23. The method of claim 11, wherein said profile attribute is
2 demographic information.

1 24. The method of claim 11, wherein said profile attribute is a gender
2 of said test user.

1 25. A machine-implemented method for extrapolating profile
2 information from web page access patterns of a test user, comprising:

3 detecting a set of web pages accessed by a test user;
4 counting said detected web pages to obtain a total number of test user web
5 pages;
6 if said total number of test user web pages is in a first range, then
7 performing the steps of:
8 initializing a first set of Expectation Maximization (EM)
9 parameters,
10 performing a first EM process using said first set of initialized
11 parameters to obtain a first EM process result, and
12 assigning a user profile attribute to said test user in response to
13 said first EM process result; and
14 if said total number of test user web pages is in a second range, said first
15 and second ranges do not overlap, then performing the steps of:
16 initializing a second set of EM parameters,
17 performing a second EM process using said second set of
18 initialized parameters to obtain a second EM process result, and
19 assigning a user profile attribute to said test user in response to
20 said second EM process result.

1 26. A machine-implemented method for extrapolating profile
2 information from web page access patterns of a user, comprising:
3 performing a vector classification method to obtain a vector classification
4 result;

5 performing a bias classification method to obtain a bias classification
6 result;
7 performing a probabilistic classification method to obtain a probabilistic
8 classification result;
9 combining at least two of said results to obtain a combination result; and
10 assigning a user profile attribute to said user in response to said
11 combination result.

1 27. A machine-implemented method for extrapolating profile
2 information from web page access patterns of a user, comprising:
3 detecting a set of web pages accessed by a user;
4 counting web pages in said set of web pages to obtain a total number of
5 web pages;
6 performing a first classification method to obtain a first classification
7 result if said total is within a first range;
8 performing a second classification method to obtain a second
9 classification result if said total is within a second range, wherein each of said
10 first and second classification methods are selected from the group consisting of
11 vector, bias, and probabilistic classification methods; and
12 assigning a user profile attribute to said user in response to at least one of
13 said results.

1 28. An apparatus for extrapolating user profile information from web
2 page access patterns of a user, comprising:

3 a memory, said memory adapted to store program code;
4 a processor in communication with said memory, said program code
5 capable of programming said processor to perform a method for extrapolating
6 user profile information from web page access patterns of a user, the method
7 comprising:
8 detecting a set of web pages accessed by a user;
9 mapping at least a subset of said web pages to a first data
10 structure, said first data structure identifies a web page access pattern of said user;
11 comparing said first data structure to a second data structure to
12 obtain a comparison result, said second data structure identifies a web page
13 access pattern of a set of known users, said known users having a user profile
14 attribute in common; and
15 assigning said user profile attribute to said user in response to said
16 comparison result.

1 29. An apparatus for extrapolating user profile information from web
2 page access patterns of a user, comprising:

3 a memory, said memory adapted to store program code;
4 a processor in communication with said memory, said program code
5 capable of programming said processor to perform a method for extrapolating
6 user profile information from web page access patterns of a user, the method
7 comprising:
8 detecting a set of web pages accessed by a test user;

9 initializing a first set of Expectation Maximization (EM)
10 parameters;
11 performing a first EM process using said first set of initialized
12 parameters to obtain a first EM process result; and
13 assigning a user profile attribute to said test user in response to
14 said first EM process result.

1 30. A processor readable storage medium, comprising:
2 processor readable program code embodied on said processor readable
3 storage medium, said processor readable program code for programming a
4 processor to perform a method for extrapolating user profile information from
5 web page access patterns of a user, the method comprising:
6 detecting a set of web pages accessed by a user;
7 mapping at least a subset of said web pages to a first data
8 structure, said first data structure identifies a web page access pattern of said user;
9 comparing said first data structure to a second data structure to
10 obtain a comparison result, said second data structure identifies a web page
11 access pattern of a set of known users, said known users having a user profile
12 attribute in common; and
13 assigning said user profile attribute to said user in response to said
14 comparison result.

1 31. A processor readable storage medium, comprising:

2 processor readable program code embodied on said processor readable
3 storage medium, said processor readable program code for programming a
4 processor to perform a method for extrapolating user profile information from
5 web page access patterns of a user, the method comprising:

6 detecting a set of web pages accessed by a test user;
7 initializing a first set of Expectation Maximization (EM)
8 parameters;

9 performing a first EM process using said first set of initialized
10 parameters to obtain a first EM process result; and

11 assigning a user profile attribute to said test user in response to
12 said first EM process result.